Empirical Tools of Public Economics

Part-1
Outline

3.1. Correlation vs. Causality
3.2. Ideal case: Randomized Trials
3.3. Reality: Observational Data
Correlation vs. Causality

• Correlation: Two economic variables are correlated if they move together.

• Causality: Two economic variables are causally related if the movement of one causes movement of the other.
Correlation vs. Causality

• Q: Which one do we need?
• A: Depends on the context:
  – To test the theoretical predictions, we need causality. (Ex: to estimate the impact of TANF on the labor market supply of single mothers.)
  – In order to forecast the future value of an economic variable, correlation is sufficient.
Correlation vs. Causality

• It is crucial to realize the difference between causality and correlation. Examples:
  – More police officers (A) in high crime areas (B)
  – My grandmother’s arthritis (A) and rain (B)
  – Taking SAT preparation course (A) and SAT scores (B)
Correlation vs. Causality

• Assume you found that event A is correlated with event B.
  – A is causing B
  – B is causing A
  – Some third factor is causing both
Correlation vs. Causality

• SAT Example:
  – Taking SAT prep courses lowers SAT scores.
  – Those who are of lower test-taking ability take preparation courses to catch up.
  – Those who are generally nervous like to take prep courses and being nervous is associated with lower SAT scores.
Correlation vs. Causality

• SAT Example:
  – Simple correlation does not tell us which one of these three possible scenarios is valid.
  – How to determine?
Ideal Case: Randomized Trials

• Typical example: How do new medical treatments affect the health of medical patients?
  – Example: the impact of sleeping pills on patients suffering insomnia
Ideal Case: Randomized Trials

- Randomized trial
  - Given a group of volunteers, randomly assign some of them to the control group and some to the treatment group.
  - Treatment group: the set of individuals who are subject to an intervention being studied.
  - Control group: the set of individuals comparable to the treatment group who are not subject to the intervention being studied.
Ideal Case: Randomized Trials

• Randomized trial-Insomnia Example
  – Given a group of volunteer insomnia patients, randomly assign some of them to the control group (not receiving the sleeping pill) and some to the treatment group (receiving the sleeping pill).
  – At the end of the experiment, compare the change in the average hours slept among the treatment group to the change among the control group.
Ideal Case: Randomized Trials

• Why do we need randomization?

  – If not, the treatment group and the control group might differ in characteristics other than the treatment that both affect the selection into treatment and the outcome of interest leading to *bias*. 
Ideal Case: Randomized Trials

• Why do we need randomization?
  – SAT Example:
    • Treatment group: Students taking SAT prep courses
    • Control group: Students not taking SAT prep courses
    • Outcome of interest: SAT scores
Ideal Case: Randomized Trials

• Why do we need randomization?
  – SAT Example:
    • Observe the course-takers and non-takers and compare their SAT scores.
    • Issue: If students with low test-taking ability take SAT courses, the observed difference between the SAT scores of the treatment group and the control group might not be due to the SAT courses.
    • The difference might be rather due to the difference in abilities.
Ideal Case: Randomized Trials

• Why don’t we always use randomized trials?
  – Ethical reasons: Proposed treatment for the Parkinson’s disease
    • Inject fetal pig cells directly into patients’ brains
      – Drilled holes in the heads of all 18 volunteers, but put the pig cells only in 10 of the subjects.
  – Attrition bias: Reduction in the size of samples over time, which, if not random, can lead to biased estimates.